



Research module project (Forschungsmodul)

Comparison of TE content and distribution between *Drosophila* species

Background: *De novo* genes are genes that emerge from scratch from noncoding regions of a genome. To do this they first need to gain both transcription and an open reading frame. Recent studies have shown that transposable elements (i.e. mobile genetic elements that can move and amplify in the genome) could play a role in *de novo* gene emergence by providing regulatory motifs that could enable novel transcription events. We aim to further study these dynamics using a dataset of genomes from six *Drosophila* species with multiple strains from different geographical populations per genome. Prior to focusing on *de novo* genes, it would be beneficial to compare TE dynamics within and between the different species.

Objectives: The student will make use of the dataset of 37 genomes and transcriptomes from the six *Drosophila* species. Additionally they will get the corresponding TE annotation of the genomes. Using this they will compare the TE dynamics of the genomes within and between the six species. Potential interesting analysis could include:

- Looking at the TE distribution across the genomes. Are different TE families distributed differently?
- Comparing TE content between the species (whole genome/CDS/introns/intergenic, ...)

Requirements:

- Experience (or interest) in using the command line
- Experience in working with R and/or python

Methods:

- Data analysis and plotting with python and/or R

Supervision: Marie Lebherz, Room 100.20, m.lebherz@uni-muenster.de, Molecular Evolution and Bioinformatics Group (<https://bornberglab.org/people/lebherz/>).

Selected Literature:

- 1) Lebherz, M. K., Fouks, B., Schmidt, J., Bornberg-Bauer, E., & Grandchamp, A. (2024). DNA Transposons Favor De Novo Transcript Emergence Through Enrichment of Transcription Factor Binding Motifs. *Genome Biology and Evolution*, 16(7), evae134.
- 2) Grandchamp, A., Kühn, L., Lebherz, M., Brüggemann, K., Parsch, J., & Bornberg-Bauer, E. (2023). Population genomics reveals mechanisms and dynamics of de novo expressed open reading frame emergence in *Drosophila melanogaster*. *Genome Research*, 33(6), 872-890.